Genetic Ensemble Biased ARTMAP Method of ECG-Based Emotion Classification

Chu Kiong Loo, Wei Shiung Liew, and M. Shohel Sayeed

Abstract. This study is an attempt to design a method for an autonomous pattern classification and recognition system for emotion recognition. The proposed system utilizes Biased ARTMAP for pattern learning and classification. The ARTMAP system is dependent on training sequence presentation to determine the effectiveness of the learning processes, as well as the strength of the biasing parameter, lambda λ. The optimal combination of λ and training sequence can be computed efficiently using a genetic permutation algorithm. The best combinations were selected to train individual ARTMAPs as voting members, and the final class predictions were determined using probabilistic ensemble voting strategy. Classification performance can be improved by implementing a reliability threshold for training data. Reliability metric for each training sample was computed from the current voter output, and unreliable training samples were excluded from the performance calculation. Individual emotional states are highly variable and are subject to evolution from personal experiences. For this reason, the above system is designed to be able to perform learning and classification in real-time to account for inter-individual and intra-individual emotional drift over time.

1 Introduction

Several methods were developed to develop emotion recognition systems based on facial and speech recognition [1] [2], as well as physiological signal measurements

Chu Kiong Loo · Wei Shiung Liew
Faculty of Computer Science and Information Technology
University of Malaya
Kuala Lumpur, Malaysia
e-mail: ckloo.um@gmail.com, liew.wei.shiung@gmail.com

M. Shohel Sayeed
Faculty of Information Science and Technology
Multimedia University
Melaka, Malaysia
e-mail: shohel.sayeed@mmu.edu.my

T. Watanabe et al. (Eds.): Intelligent Interactive Multimedia: Systems & Services, SIST 14, pp. 299–306.
springerlink.com

[3] [4] [5]. The main advantage of physiological measurements over facial and speech recognition is the practicality of implementing the monitoring systems. Physiological signal monitoring have the benefit of a constant and robust signal recording, especially with the development of portable biosensors that are also unobtrusive for daily activities.

A study by Zhong et.al [6] analyzes the nonlinear components of heart-rate dynamics caused by the two main branches of the autonomic nervous system (ANS). In addition to another study [7], these methods can be used to examine the fluctuations of the ANS. Principal dynamic modes (PDM) allows for nonlinear analysis on the separated dynamics in the ECG signal, as well as clear separation between the contributions of the two ANS branches.

Individual emotion states are subject to variations due to external and internal influences. This emotional drift requires that any autonomous emotion classification system to be regularly updated with its users' current physiological-emotion data. The system must be capable of incorporating new learning patterns while retaining previous knowledge without performing the entire learning sequence. This "stability vs. plasticity" dilemma can be minimized by using adaptive resonance theory (ART) for pattern learning and classification.

ART-based neural networks were developed as a model of human cognitive information processing. During learning or training, certain input sequences with a specific featural attention can distort the system's memory and reduce its classification accuracy. Biased ARTMAP [9] solves the problem of overemphasis on early critical features by biasing attention away from previously attended features when the system makes a predictive error. The strength of the biasing is controlled by an attention parameter, λ.

Using Biased ARTMAP for pattern recognition reduces the number of variables which determines the system's performance to two factors: the attention parameter λ, and the sequence of the training data presentation. Optimum combinations of λ and training data sequence can be computed efficiently by implementing a genetic permutation method [10] to "evolve" the optimal combinations over several generations.

To further improve the classification accuracy, a voting strategy is used to determine the final class predictions. The proposed voting strategy [11] calculates recognition rates of plurality voting techniques while considering the system's measure of reliability, that is, the probability of a decision to be classified correctly given a specific input pattern. Using the reliability metric to describe a given training data can reduce the classification error by rejecting suspicious training data which do not meet a minimum reliability requirement [12].

The final incarnation of the classification system will be a prototype emotion recognition system that can be customized for individual by continuous feedback of ECG measurements into the system to improve predictive accuracy of the individual's emotions. Constant online learning will generate enough data for the system to adapt to its user's emotional drift.

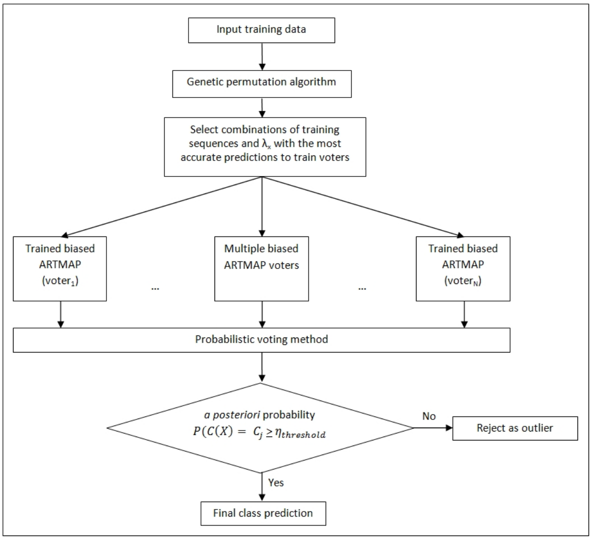

Fig. 1.1 Block diagram of the Genetic Ensemble Biased ARTMAP system

2 Genetic Ensemble Biased ARTMAP

2.1 Genetic Permutation for Optimal Pattern Ordering

20 chromosomes were generated for the initial population. Each chromosome consists of a number of genes equal to the number of feature patterns, ***N***, in each training sample. Thus, the gene arrangement for each chromosome is representative of a single data presentation sequence of the given data set.

To calculate the fitness value of each chromosome, a single-voter Biased ARTMAP was trained and tested with 5-fold cross-validation. Fitness value of the chromosome was calculated as the percentage of correctly classified patterns over the total number of tested patterns. Chromosomes were then sorted according to fitness, and half of the least fit chromosomes were discarded.

Reproduction was performed to repopulate the chromosome pool and replace the discarded chromosomes with offspring of randomly-chosen fit parents. Mutation was also applied to randomly selected offspring to prevent early convergence. The process of fitness testing and selection, mating, and mutation ensured that each successive generated population of chromosomes will have a higher average fitness than the previous generation.

2.2 Biased ARTMAP Theory

ART-based neural networks model real-time prediction, searching, learning, and recognition, capable of learning stable recognition categories in response to arbitrary input sequences. One example of the ART learning system is the ARTMAP, a hierarchical network architecture that can self-organize stable categorical mappings between m-dimensional input vectors and n-dimensional output vectors [13].

The novelty of the ARTMAP's learning strategy on attended critical feature patterns possesses a design flaw which presents itself during online fast learning. Extensive testing show that certain input sequences with a strong featural attention can distort the learning process and affect the system's testing accuracy.

This flaw is addressed by adding a module to the ARTMAP system to selectively bias previously attended features whenever the system makes a predictive error. Biasing the input pattern to favour previously inactive nodes implements search by allowing the network to activate new nodes, and biases the system against reselecting the category node which had just produced the predictive error.

2.3 Probabilistic Voting

The probabilistic voting strategy was proposed as an alternative to majority voting [11]. In general, the error rate of the combination system was minimized by choosing the classification category with the largest *a posteriori* probability according to Bayes' rule, while all other classes have equal probability of being chosen in case of incorrect classification. Each classifier can have a different weight and each class has a constant representing its *a priori* probability.

A study by Loo and Rao [12] implements a method in extension of the above to measure the reliability of a class prediction computed from the probabilistic voting results. The desired reliability of a classification system can be enforced by requiring that each and every input object's winning class to have at least ***r*** more votes than the closest competing class, failure of which the classification of the input object is rejected due to unreliability of the prediction.

3 Experiment

The performance of the Genetic Ensemble ARTMAP system was tested using several datasets from the UCI Machine Learning Repository [15]. The tested datasets were Dermatology, Glass, Hepato, and Wine, chosen for non-binary classification. Optimization was performed for each data set using the same methods outlined above. The resultant pool of 220 potential training sequences were then used for training a single-voter, five-voter, and ten-voter system. In addition, the 220 training sequences were used to generate a bootstrapped mean with 1000 resamplings and 95% confidence interval, as a representation of the overall classification accuracy of each data set.

Table 3.1 Prediction accuracy of bootstrapped mean of the genetic optimized population, and the probabilistic voting system

	Bootstrapped population			Probabilistic ensemble voting		
	Low	Mean	High	1-voter	5-voters	10-voters
Derm	93.84	94.00	94.16	97.74	96.05	96.05
Glass	67.85	68.48	69.07	82.38	82.38	76.19
Hepato	70.05	70.17	70.29	72.52	70.65	69.90
Wine	89.56	89.71	89.86	92.57	90.28	89.71

The obtained results, when compared to similar literature using other pattern classification methods with the same data sets, show comparable performance of the Genetic Ensemble Biased ARTMAP with other contemporary methods.

The system is then tested using a database of physiological signals collected by Wagner et.al. [3]. The data set consists of 100 ECG samples divided into four emotion classes. Feature extraction was performed using an algorithm designed by Wagner et.al. [14]. A total of 106 ECG features were extracted from each recording, including several features not available in the original toolbox algorithm. The features of principal dynamic modes [6] [7] were included to provide nonlinear analysis to the overall feature set.

A genetic optimization algorithm was employed to obtain optimal training sequences for the data set for each value of λ from 0 to 10. The genetic optimization process was iterated for 20 generations, and then repeated with another random population of chromosomes for a different value of λ. A total of 220 chromosomes were generated from the optimization exercise and the chromosomes with the best fitness were chosen to train a Biased ARTMAP each. A probabilistic voting strategy was then used to determine the final class prediction for any given data input.

For each value of λ, 220 training sequences were generated using the genetic optimization algorithm. The predictive accuracy of each individual training sequences were obtained, and were used to generate a bootstrapped mean as a statistical aggregate of the entire population's predictive accuracy. Results show little distinction in predictive accuracy when different values of λ were used. All bootstrapped mean results were clustered around 65-67% accuracy, while the individual predictive accuracies range from 54-78% prediction rate.

One hypothesis is that genetic ordering compensation inadvertently solves the problem of early featural distortion which the Biased ARTMAP was designed to solve. Nevertheless, the results were obtained from offline learning, and the biasing technique will be more useful during online learning.

A probabilistic ensemble voting system was applied, in which N voters were individually trained by N of the best training sequences from the combined population of 220 chromosomes. Testing was performed on the voting system based on probabilistic majority rules to determine the final class prediction of the test data. Testing was repeated using a reliability metric to evaluate each class prediction. Class predictions which did not meet the reliability threshold were removed from the final accuracy calculation.

Table 3.2 Classification performance of Fuzzy ARTMAP ($\lambda = 0$) with probabilistic ensemble voting and reliability threshold

Voters	Reliability R=0	R=0.5	R=0.9	R=0.99
1	76.00 (0)	76.00 (0)	NaN (100)	NaN (100)
3	76.00 (0)	76.76 (1)	85.13 (26)	85.13 (26)
5	70.00 (0)	70.40 (2)	77.64 (15)	78.31 (17)
7	71.00 (0)	71.42 (2)	76.13 (12)	78.57 (16)
10	73.00 (0)	74.22 (3)	73.95 (4)	73.33 (10)

Table 3.3 Classification performance of Biased ARTMAP (λ = 0:10) with probabilistic ensemble voting and reliability threshold

Voters	Reliability R=0	R=0.5	R=0.9	R=0.99
1	78.00 (0)	78.00 (0)	NaN (100)	NaN (100)
3	79.00 (0)	79.00 (0)	87.17 (22)	87.17 (22)
5	79.00 (0)	79.38 (3)	84.44 (10)	84.88 (14)
7	75.00 (0)	74.48 (2)	77.77 (10)	78.65 (11)
10	79.00 (0)	79.78 (6)	80.64 (7)	80.89 (11)

The number in brackets represents the percentage of class predictions which were rejected due to low reliability. In particular, predictions from a voting system with few voting members are considered less reliable due to lack of information compared with systems with more voting members. However, this experiment also indicates that while predictive accuracy increased when a more stringent reliability threshold was applied, increasing the number of voters did not elicit an improvement. This may be explained by the method in which each voter was trained. Each additional voter besides the first was trained using training sequences which were increasingly less accurate, effectively affecting the system's predictive accuracy by adding an increasing amount of noisy data. Even so, each additional voter served to contribute additional information into the ensemble classifier by improving recognition rates of reliable training data.

The above results were then compared against similar pattern classification methods: linear discriminant analysis (LDA), k-nearest neighbor (kNN), and multilayer perceptron (MLP). For kNN, a series of training and testing was performed for a range of values for k, in the range [1, 10]. A bootstrapped mean was generated from the results. For MLP, the main initial network parameters are the number of hidden layers (set to 9), the rate of learning (set to 1), and the number of training iterations (set to 100). For Fuzzy ARTMAP and Biased ARTMAP, the results were using the classification performance from the best combination of voter ensemble, reliability threshold, and training sequence.

Table 3.4 Comparison of pattern classification methods

Classification method	Predictive accuracy (%)
Linear discriminant analysis	66.00
K^{th} nearest neighbor [k = 1:10]	72.00
Multilayer perceptron	83.00
Genetic Ensemble Fuzzy ARTMAP	85.13 (3-voter, 90% reliability)
Genetic Ensemble Biased ARTMAP	87.17 (3-voter, 90% reliability)

Both ARTMAPs show comparable classification performance with the multilayer perceptron (MLP). However, ARTMAPs have several distinct advantages over the MLP classification method, including the ability for incremental learning to evolve the classification system over time, and a faster convergence during training and testing.

Conclusion

From the experiment, several conclusions can be drawn. The genetic optimization algorithm is an effective method for training and testing an ARTMAP system for pattern learning and classification. When combined with Biased ARTMAP, the genetic optimization method rendered the biasing technique redundant. In addition, a more effective voter selection method will be required, as the current method reduces the predictive accuracy of the system with each additional voter added. Implementing a reliability threshold allows a slight increase in classification accuracy by rejecting class predictions which do not meet the minimum consensus among voting members, as opposed to simple majority voting. Overall, the Genetic Ensemble Biased ARTMAP has comparable pattern prediction capability as compared with other pattern classification methods such as LDA, kNN, and MLP. The system's features can be further improved based on the results from this study.

References

[1] De Silva, L.C., Miyasato, T., Nakatsu, R.: Facial emotion recognition using multimodal information. In: Proceedings on International Conference on Information, Communications, and Signal Processing, vol. 1, pp. 397–401 (1997)

[2] Busso, C., Deng, Z., Yildrim, S., Bulut, M., Lee, C.M., Kazemzadeh, A., Lee, S., Neumann, U., Narayanan, S.: Analysis of emotion recognition using facial expressions and multimodal information. In: Proceedings of the 6th International Conference on Multimodal Interfaces, pp. 205–211 (2004)

[3] Wagner, J., Kim, J., Andre, E.: From physiological signals to emotion: Implementing and comparing selected methods for feature extraction and classification. In: IEEE International Conference on Multimedia and Expo, pp. 940–943 (2005)

[4] Kim, K.H., Bang, S.W., Kim, S.R.: Emotion recognition system using short-term monitoring of physiological signals. Medical and Biological Engineering and Computing 42(3), 419–427 (2004)

[5] Mandryk, R.L., Atkins, M.S.: A fuzzy physiological approach for continuously modeling emotion during interaction with play technologies. International Journal of Human-Computer Studies 65(4), 329–347 (2007)

[6] Zhong, Y., Wang, H., Ju, K.H., Jan, K.M., Chon, K.H.: Nonlinear analysis of the separate contributions of autonomics nervous systems to heart rate variability using principal dynamic modes. IEEE Transactions on Biomedical Engineering 51(2), 255–262 (2004)

[7] Choi, J., Gutierrez-Osuna, R.: Using heart rate monitors to detect mental stress. In: 6th International Workshop on Wearable and Implantable Body Sensor Networks, pp. 219–223 (2009)

[8] Plutchik, R.: The nature of emotions. American Scientist (2001)

[9] Carpenter, G.A., Gaddam, S.C.: Biased ART: A neural architecture that shifts attention toward previously disregarded features following an incorrect prediction. Neural Networks 23(3), 435–451 (2010)

[10] Palaniappan, R., Eswaran, C.: Using genetic algorithm to select the presentation order of training patterns that improves simplified fuzzy ARTMAP classification performance. Applied Soft Computing 9(1), 100–106 (2009)
[11] Lin, X., Yacoub, S., Burns, J., Simske, S.: Performance analysis of pattern classifier combination by plurality voting. Pattern Recognition Letters 24, 1959–1969 (2003)
[12] Loo, C.K., Rao, M.V.C.: Accurate and reliable diagnosis and classification using probabilistic ensemble simplified fuzzy ARTMAP. IEEE Transactions on Knowledge and Data Engineering 17(11), 1589–1593 (2005)
[13] Carpenter, G.A., Grossberg, S., Reynolds, J.H.: ARTMAP: A self-organizing neural network architecture for fast supervised learning and pattern recognition. Neural Networks 4(5), 565–588 (1991)
[14] Wagner, J.: The Augsburg Biosignal Toolbox (2009), `http://www.informatik.uni-augsburg.de/en/chairs/hcm/projects/aubt/` (retrieved June 29, 2011
[15] Frank, A., Suncion, A.: UCI Machine Learning Repository Irvine, CA: University of California, School of Information and Computer Science (2010), `http://archive.ics.uci.edu/ml` (retrieved November 2011)